



Degree project in Computer Science

First cycle, 15 credits

Diminishing returns of higher mp3 bit rates

A literature review

MARTIN BÖRJESON, ERIK PERSSON

Diminishing returns of higher mp3 bit rates: a literature review

MARTIN BÖRJESON

ERIK PERSSON

Degree Project in Computer Science and Engineering

Date: June 30, 2022

Supervisor: Jeanette Hällgren Kotaleski

Examiner: Pawel Herman

School of Electrical Engineering and Computer Science

Swedish title: Icke-märkbarhet av högre mp3-bitrates: en
literaturstudie

Abstract

Digital audio is the by far most common way to listen to music today. It is therefore important to store audio files as efficiently as possible. A common way of doing this is by using compression, often with mp3 files. As with all non-lossless compression, some data is lost. The purpose of this literature review is to examine at what point the compression is indiscernible. Four research papers were studied and their results compared and compiled to draw a more broad-based conclusion to the research question. The papers all used A/B tests to test at what bit rate subjects could not tell the difference between a compressed and an uncompressed file. The results indicate that most people with consumer-grade equipment are incapable of discerning between mp3 files encoded at 128 kbit/s mp3 and uncompressed or lossless audio. With trained ears and high-end equipment, the cutoff bit rate is instead in the 256 kbit/s range.

Sammanfattning

Nästan all musik som lyssnas på idag är digital. Därför är det viktigt att kunna lagra ljudfiler så effektivt som möjligt. Ett vanligt sätt att göra detta är genom att använda komprimering, ofta med mp3-filer. Mp3 och många andra effektiva komprimeringsmetoder sker dock med viss degredering av datan. Syftet med denna litteraturstudie är att undersöka när komprimeringen av mp3-filer blir omärkbar. Fyra forskningsartiklar valdes ut och studerades. Deras resultat jämfördes och samlades för att kunna dra en mer slutsats med bredare bas. Artiklarna använde alla A/B-tester för att undersöka vid exakt vilken bitrate kvalitétsskillnaderna mellan en komprimerad mp3-fil och okomprimerad ljudfiler slutade vara märkbara. Resultaten tyder på att de flesta med konsumentutrustning inte hör någon skillnad på 128 kbit/s mp3-filer och förlustfria ljudfiler. Med träning och professionell utrustning är det dock möjligt att uppfatta skillnader på upp till 256 kbit/s.

Contents

1	Introduction	1
1.1	Research Question	1
2	Background	2
2.1	Lossy vs Lossless compression	2
2.2	Bit rate	3
2.3	Mp3	3
2.4	Methods for measuring perceived audio quality	4
2.5	Bit rates of various streaming services	5
3	Methods	6
4	Results	8
4.1	Pras <i>et al.</i> (2009)	8
4.2	Ruzanski (2006)	9
4.3	Lee, Horner (2010)	10
4.4	Böhne <i>et al.</i> (2013)	10
5	Discussion	12
5.1	Method discussion	12
5.2	Result discussion	13
6	Conclusions	15
7	Future studies	16
	Bibliography	17

Chapter 1

Introduction

A recent trend among some audio streaming services is to provide higher quality audio streams. Tidal was launched 7 years ago, with the premise of bringing high quality audio streaming to the masses [1]. In June 2021, Apple began offering lossless audio as a free upgrade on their own music streaming platform, Apple Music [2]. This indicates a clear demand of higher quality audio from customers, but can most people even tell the difference between low and high quality audio? If not, then the only differentiating factor of a hi-fi streaming service might be its higher price and data rate. It has been shown that people enjoy a wine more if they think it is expensive [3], might this be another case of a similar phenomenon?

1.1 Research Question

Our research question is thus as follows: *At what bit rate are mp3 files generally perceived as lower quality than CD? (And how does this compare to the commonly used bit rates in the music streaming industry?)*

A necessary assumption to answer the question is that an mp3 file with a sufficiently high bit rate is indistinguishable from a lossless audio file.

Chapter 2

Background

2.1 Lossy vs Lossless compression

Most commonly used file formats compress the audio data to reduce file size. This is achieved by encoding the original data in a way that decreases the statistical redundancy of the information. Patterns and repetition in the data can often be described in a more compact way. Data without patterns of any kind, i.e. random data, has no statistical redundancy, and cannot be appreciably compressed.

Data compression can be done with or without data loss (lossy or lossless). The goal of lossy compression is to throw away unimportant detail in the data while preserving the important parts. Figures 2.1 and 2.2 show how lossy image compression might make areas with many colors of similar shades into areas of a singular color, while still obviously being from the same image.

Lossless compression preserves all information, which means that the original data can be reconstructed without flaws from a losslessly compressed file. For some kinds of files, such as text documents, it is paramount that not a single character is altered in the process of compression. Such files are therefore always compressed using a lossless algorithm (e.g. .ZIP archive files). This is in stark contrast to the compression of video and audio, where many details can be removed with no perceptible difference to the uncompressed original. The data compression ratio, i.e. the size of the uncompressed file divided by the size of the compressed file, is almost always higher for lossy algorithms compared to lossless algorithms.



Figure 2.1:
A less compressed JPEG



Figure 2.2:
A highly compressed JPEG

2.2 Bit rate

In audio coding, the bit rate is the amount of bits required to encode a second of audio. In the uncompressed CD audio format, it is the product of the amount of samples per second and the amount of bits used to represent each sample, and is often used as a broad indicator of audio quality. CD quality usually implies a sample size of 16 bits, and a sample rate of 44100 Hz. Multiplying these values together yields a bit rate of 705.6 kbit/s per channel (i.e left and right). The available bit rates in the mp3 standard are all significantly lower than the uncompressed CD quality, and the data thus has to be compressed.

2.3 Mp3

Mp3 is a widely used lossy audio coding format. There are several different formats of mp3 files, categorized under MPEG-1 Audio Layer III, MPEG-2 Audio Layer III and MPEG-2.5 Audio Layer III. These different formats have several different bit rates, detailed in table 2.1. It is able to achieve a high level of compression by reducing the quality of the audio components which are considered to be beyond the capabilities of most human hearing, and preserving fidelity in the more noticeable areas. Originally released in 1991 [4], the format has remained prevalent in the present. After the last mp3-related patent expired in 2017, mp3 became an open and free format.

MPEG-1 Audio Layer III	MPEG-2 Audio Layer III	MPEG-2.5 Audio Layer III
–	8	8
–	16	16
–	24	24
32	32	32
40	40	40
48	48	48
56	56	56
64	64	64
80	80	–
96	96	–
112	112	–
128	128	–
n/a	144	–
160	160	–
192	–	–
224	–	–
256	–	–
320	–	–

Table 2.1: Different bit rates (kbit/s) available to different mp3 formats

There are several commonly used encoders for mp3 files, with one of the most popular being LAME.

2.4 Methods for measuring perceived audio quality

There are many standards for measuring perceived audio quality, and it would take a research paper in its own right to explore and compare their respective differences, advantages and disadvantages. Many of the standards do however rely on the same underlying technique, namely a blind (or double blind) A/B test, where the subjects differentiate between two recordings which differ in some minor way (audio quality in the case of this study). It is crucial that the comparisons really are blind, as the subjects might be biased towards selecting the sample which they are told is higher quality.

2.5 Bit rates of various streaming services

In table 2.2, the minimum and maximum used bit rates of a few popular streaming services are shown. Note that the maximum bit rate might only be available in a higher paid tier of the service. The minimum bit rate is in some cases lower than the default bit rate of the service.

Service	Minimum Bit rate (kbit/s)	Maximum Bit rate (kbit/s)
Spotify [5]	24	320
Apple Music [6]	not specified	9216 (lossless)
Tidal [7]	160	9216 (lossless)
Youtube Music [8]	48	256
SoundCloud [9]	64	256

Table 2.2: Bit rates of a selection of popular streaming services

Chapter 3

Methods

To ascertain the cutoff point, a literature review was made. A literature review was chosen because it allowed research without using expensive audio equipment which was not available to the authors at the time of writing. Additionally, no test subjects were required.

This literature review used quantitative sources, as perceived audio quality at various bit rates can easily be measured. The aim was to find commonalities across sources, and to determine at which bit rates the audio quality of mp3s generally are indistinguishable from uncompressed or lossless formats. The lowest bit rate in each study at which the difference in preference was deemed to be insignificant ($p > 0.05$) was selected. The p-value was calculated with a binomial test with the probability of correctly identifying a track set to 50%. This probability was chosen due to it signaling an inability to discriminate the lower bit rate sound, and instead choosing randomly, i.e. with a 50/50 chance to choose either version.

To find relevant sources, multiple methods were used. The first sources were found using Google Scholar, searched with the keywords ‘audio quality’, ‘perceived audio quality’ and ‘audio quality mp3’. This led to several relevant studies being found. Additionally, the cited sources in these studies was investigated for additional relevant material. After some relevant studies were found, a search was performed (via ResearchGate) to find other studies that cited the original studies, and these were also investigated and analyzed.

As the studies’ methodology varies, their results and conclusions might not be directly comparable. This is a common complication with many literary studies. To minimize the issue, only sources which utilized a double blind

A/B test were considered for the study.

In the end, the following four papers matched all criteria:

Title	Author	Publication
Subjective Evaluation of mp3 Compression for Different Musical Genres [10]	Pras, Zimmerman, Levitin and Guastavino	Proceedings of the 127th Convention of the Audio Engineering Society
Effects of mp3 encoding on the sounds of music [11]	Ruzanski	IEEE Potentials 25(2):43 - 45 (doi: 10.1109/MP.2006.1649011)
Discrimination of mp3-Compressed Musical Instrument Tones [12]	Lee, Horner	The Journal of the Acoustical Society of America 128(4):2284 (doi: 10.1121/1.3508003)
Subjective audibility of mp3-compression artefacts in practical application [13]	Böhne, Gröger, Hammerschmidt, Helm, Hoga, Kraus, Rösch, Sussek	Student work at Hamburg University

Table 3.1: Examined papers

Expected and unexpected discrepancies between sources are to be reviewed in the discussion.

Chapter 4

Results

In this chapter, each paper is described, including their method as well as results. In table 4.1 the cutoff point, where higher bit rates are indistinguishable from lower ones, for each study is shown.

Author(s)	Bit rate (kbit/s)
Pras et al.	256
Ruzanski	128
Lee, Horner	128
Böhne et al.	96

Table 4.1: Cutoff point for each paper

4.1 *Pras et al.* (2009)

At the Audio Engineering Society's 127th convention, *Pras et al.* presented their research paper about how participants with different areas of expertise subjectively evaluated the audio quality of mp3 music, separated by genre [10]. The procedure involved conducting double blind comparative A/B tests and a three-part questionnaire. Subjects would get presented with two audio files of the same song, with different bit rate encoding (two of 96, 128, 192, 256 and 320 kbit/s), encoded with LAME, and had to choose which file they preferred in terms of quality.

The authors found that for mp3 files with bit rates between 96 and 192 kbit/s, listeners would consistently prefer a higher quality file. Audio technicians

would more consistently identify the higher quality file than musicians. The researchers found this to be a plausible result, as audio technicians would have more experience recognizing the specific sound artifacts incurred by a lower bit rate encoding. It was suggested that musicians may not focus on the same sound criteria. Musicians were also observed to perform significantly better on the tests with excerpts which included their instruments of choice. The listeners could more easily pick up artifacts in electronic tracks compared to acoustic tracks, which the researchers believed some might find counter-intuitive. The explanation they found was that mp3-encoding was generally believed to require more headroom (i.e. difference between the maximum quantifiable sound level and the peak level of the audio signal) than CD encoding.

The audio equipment used was a Grace m906 DAC and monitor system, Classé CA-5200 stereo amplifier, and a pair of Bowers & Wilkin 902 Diamond loudspeakers. The subjects consisted of thirteen participants, of which nine were audio technicians and four were musicians.

Since Pras *et al.* already used a binomial test, one was not performed by the authors of this paper. The cutoff point was determined to be 256 kbit/s.

4.2 Ruzanski (2006)

Ruzanski's article was published in *IEEE Potentials* in April 2006. In it was described a study where a blind A/B test was performed in order to find out subjects' preferences in regards to mp3 bit rate compared to CD quality. Five to ten second long pieces of music from five different genres were chosen for the tests. The bit rates that were included in the test were 32, 64, 96, 128 and 192 kbit/s. Sony MDR CD250 headphones were used to perform the tests.

The results were shown to differ quite considerably depending on the genre of music played. For instance, subjects were unable to differentiate a nocturne song encoded at 96 kbit/s with the same snippet played at CD quality, while they were able to choose the higher quality rap song over one encoded at 96 kbit/s with 90% accuracy. The author suggests that this could mean that songs with low dynamic range could be encoded with a lower bit rate without perceivable effect.

The subjects consisted of five untrained participants with no hearing impairments.

The binomial test described in the method was used to determine the cutoff

point to 128 kbit/s in this study.

4.3 Lee, Horner (2010)

The concept of different sounds being more susceptible to mp3 compression is further explored in Lee and Horner's report about discrimination of mp3-compressed musical tones. In the study, eight different instrumental tones, one tone from each instrument were compressed to ten different bit rates (32, 40, 48, 56, 64, 80, 96, 112, 128 and 160 kbit/s), using LAME. Thirty-one test subjects were given two pairs of sounds, and had to choose the pair they thought included a lower bit rate sound. The test subjects all had musical experience, with an average of 10 (and a minimum of 6) years of experience playing a musical instrument.

Their results showed that for bit rates close to and above 80 kbit/s, participants could in general not discriminate between the lower and higher bit rate tones. Results differed slightly for different instruments; participants had around a 65% probability to pick the higher bit rate sound when listening to the clarinet at 80 kbit/s, while the Oboe at the same bit rate was only chosen correctly 45% of the time.

The subjects consisted of thirty-one undergraduate students at the Hong Kong University of Science and Technology with no hearing impairments.

A literal binomial test was not done in the study, but the first result which was not in the 95% confidence interval was discarded, and this is equivalent to a binomial test. Therefore the cutoff point was determined to be 128 kbit/s.

4.4 Böhne *et al.* (2013)

The student work done by Böhne *et al.* at the Hamburg University Institute of Musicology concludes that the cutoff between discernible and indiscernible difference in sound quality is at 96 kbit/s, which is a lower bit rate than some of the other articles explored in this paper. The participating listener subjects were all musicology students, and were experienced in the digital music medium. They were tasked with performing a blind A/B test on pairs of LAME-compressed mp3 files of varying quality, or one compressed mp3 file and one uncompressed, CD quality, WAV file. The test consisted of 12 recordings across 4 musical genres: Electro, Rock, Classical, and Jazz.

For the 48 kbit/s samples, almost all participants were easily able to discern the difference between the mp3 and WAV audio files. The opposite held true for 128 kbit/s samples. Very few participants noticed any difference between the recordings, and the few that claimed they did were not consistently preferring either the mp3 or WAV audio samples. The results from the 96 kbit/s samples were as expected somewhere in between the results from 48 kbit/s and 128 kbit/s. The authors concluded their study by suggesting that further research should be put in exploring whether people that were used to listening to lower quality audio (e.g. mainly listening to music through online streaming) would prefer lower mp3 bit rates to higher ones, since that is what they are used to hearing.

The subjects consisted of twenty-one undergraduate and postgraduate students at the University of Hamburg, of varying experience with digital audio.

The binomial test described in the method was used to determine the cutoff point to 96 kbit/s in this study.

Chapter 5

Discussion

In this chapter, potential problems and improvements with the method will be discussed. Additionally, the results and why they look like they do, will also be discussed.

5.1 Method discussion

One significant issue with the method is that only four papers were taken into account when doing the literature review. This rather low amount is probably due to there not existing much research in the subject, and this is something that could be improved in the future. Another explanation is that the authors of this paper were unable to discover other studies which fit the criteria described in the method, perhaps because access they were only accessible on other websites than the ones used (ResearchGate and Google Scholar). The conclusions reached will thus have some amount of uncertainty.

In addition to the above, a practical study (i.e. conducting AB tests on test subjects) could have been performed to supplement the findings. The reason this was not done was because of time constraints, but a study to compare with would probably have improved the reliability of the results further.

The practical study, if one was conducted, would be constructed to be as similar as possible to the tests conducted by the studies examined in this paper.

Specifically, it would entail a double blind A/B test. The test subjects would be tasked to, from a pair of audio tracks, pick the one which they perceived to be of higher quality. In each pair of audio tracks, one would be compressed to

a specific constant bit rate, whereas the other would either be uncompressed or losslessly compressed. Each audio track would be compressed to the bit rates 64 kbit/s, 96 kbit/s, 128 kbit/s, 160 kbit/s, 192 kbit/s, and 256 kbit/s which would be compared to the original CD quality track. These pairs of tracks would be played to the listener in a random order. There would not be an option to pick neither track when they appear to be indistinguishable, as this was not an option in the other studies. The binomial test would instead be used to infer the point where the tracks are perceived to be of equal quality.

5.2 Result discussion

The bit rate cutoff between perceptible and imperceptible varies across our sources, ranging from 96 kbit/s to 256 kbit/s. Many of these discrepancies can likely be explained by differences in methodology, audio equipment, and subject experience with the audio medium, and this will be explored in further detail below. A notable caveat when comparing our findings to the commonly used bitrates of streaming services is that these services might use different coding standards, and might therefore perform differently from mp3s encoded at the same bit rate.

Pras *et al.*'s study consisted of thirteen participants. The researchers split their results further by participant vocation. It seems likely that these subgroups are too small to provide results of any significance. Using only four musicians as test subjects make it difficult to draw any conclusions regarding musicians' ability to perceive differences in audio quality. While the finding that musicians performed better when their own instrument was prominent in the audio tracks is intriguing, further research is needed to confirm that specific result. The listeners were able to distinguish between lossless and mp3 audio up to a bit rate of 256 kbit/s, which is the highest of all studied papers. This discrepancy can at least partially be explained by the study's use of high-fidelity audio equipment and test subjects with extensive experience with digital audio.

Ruzanski's study was very small, which is noticeable in their results. The probability of misclassification sometimes decreases with higher bit rates, which most likely is a result of noisy data. By averaging the results over the five genres, however, this noise is minimized. The result of a cutoff bit rate at 128 kbit/s, which is the same as with Lee and Horner, and close to Böhne *et al.*

Lee and Horner's report is somewhat disparate to the other studies, in that it only studies individual tones rather than full or portions of songs. This did not

seem to impact results very heavily however, with results being relatively in-line with the other papers considered. However, this might just be coincidence. The study also had thirty-one participants, which should eliminate any errors arising from chance.

Chapter 6

Conclusions

Our results indicate that, with consumer-grade audio equipment, the bit rates at which most popular streaming services operate render their audio tracks practically indistinguishable from an uncompressed version, and thus the quality offered by most music streaming services is adequate. With higher end equipment and trained ears, however, there is a perceptible improvement.

Chapter 7

Future studies

While mp3-files are still widely used in digital audio, the audio format is in some aspects surpassed by newer coding standards such as AAC. It might therefore be relevant to conduct similar a similar study such as this one, but with a more modern format.

It might also be of interest to further explore the interesting phenomenon that musicians might be better at distinguishing the audio quality of musical scores which contain the instruments they themselves are used to play.

Bibliography

- [1] Matthew Sparkes. *Tidal launches lossless music streaming in UK and US*. <https://www.telegraph.co.uk/technology/news/11192375/Tidal-launches-lossless-music-streaming-in-UK-and-US.html>. [Online; accessed 10-March-2022]. Oct. 2014.
- [2] Apple. *Apple Music announces spatial audio and lossless audio*. <https://nr.apple.com/dm4z5n0x8K>. [Online; accessed 10-March-2022]. Mar. 2022.
- [3] Hilke Plassmann et al. “Marketing actions can modulate neural representations of experienced pleasantness”. In: *Proceedings of the National Academy of Sciences* 105.3 (2008), pp. 1050–1054. doi: 10.1073/pnas.0706929105. eprint: <https://www.pnas.org/doi/pdf/10.1073/pnas.0706929105>. URL: <https://www.pnas.org/doi/abs/10.1073/pnas.0706929105>.
- [4] Ketan Patel, Brian C. Smith, and Lawrence A. Rowe. “Performance of a Software MPEG Video Decoder”. In: *Proceedings of the First ACM International Conference on Multimedia*. MULTIMEDIA '93. Anaheim, California, USA: Association for Computing Machinery, 1993, pp. 75–82. ISBN: 0897915968. doi: 10.1145/166266.166274. URL: <https://doi.org/10.1145/166266.166274>.
- [5] Spotify. *Audio Quality*. <https://support.spotify.com/us/article/audio-quality/>. [Online; accessed 11-April-2022]. 2022.
- [6] Apple. *About lossless audio in Apple Music*. <https://support.apple.com/en-us/HT212183>. [Online; accessed 11-April-2022]. 2021.

- [7] Tidal. *Clearly the best sound*. <https://tidal.com/sound-quality>. [Online; accessed 11-April-2022].
- [8] Google. *Select or Change your Audio Quality*. <https://support.google.com/youtubemusic/answer/9076559?hl=en>. [Online; accessed 11-April-2022].
- [9] MusConv. *Is SoundCloud sound quality better than YouTube?* <https://musconv.com/soundcloud-sound-quality-vs-youtube/>. [Online; accessed 11-April-2022].
- [10] Amandine Pras et al. “Subjective Evaluation of MP3 Compression for Different Musical Genres”. In: Oct. 2009, pp. 9–12.
- [11] Evan Ruzanski. “Effects of MP3 encoding on the sounds of music”. In: *Potentials, IEEE* 25 (Apr. 2006), pp. 43–45. doi: 10.1109/MP.2006.1649011.
- [12] Chung Lee and Andrew Horner. “Discrimination of MP3-Compressed Musical Instrument Tones”. In: *The Journal of the Acoustical Society of America* 128 (Oct. 2010), p. 2284. doi: 10.1121/1.3508003.
- [13] Hendrik Böhne et al. “Subjective audibility of MP3-compression artefacts in practical application”. In: (Sept. 2013).

